A TYPE II RESTRICTION ENDONUCLEASE AND APPLICATION THEREOF

BACKGROUND OF THE INVENTION

1.Field of the Invention

[0001] The present invention relates to a DNA-cutting enzyme, especially relates to a novel type II restriction endonuclease which recognizes and cuts DNA only at a particular sequence of nucleotides.

2. The Prior Arts

[0002] Restriction endonclease is one of DNA-cutting enzymes found in bacteria. A restriction enzyme recognition sequence containing a two fold axis of symmetry. Therefore the recognition sequences from the 5'-end to the 3'-end are the same on either upper strand or lower strand of DNA duplex, and such symmetry is termed palindrome. For the nomenclature of restriction enzymes, the first 3 letters of the name refer to a strain of bacterium, which bacterium is the source of the enzyme, the following letter indicates the particular strain, and the last part of the name is a Roman numeral which indicates the order of discovery. For example, EcoRI was isolated from *Escherichia coli* (strain RY13).

[0003] Traditionally, the restriction endonucleases are divided into 3 groups, designated type I, type II and type III according to domain structure, cleavage position, specificity of recognition sequence, and cofactors requirements. Type I and type II enzymes are similar in that both have restriction endonuclease and methylase activities. Type I restriction enzymes bind to the recognition site and then cut randomly, somewhere thousand of bases from the recognition sequence. Type III

enzymes cleave DNA around 24 to 26 bases along the length of the molecule. Type II restriction enzymes are found in late 1960's by Hamilton Smith et al., which recognize and cut DNA only at a particular sequence of nucleotides. Generally speaking, type II restriction enzymes recognize a specific sequence with 4 to 8 base pairs in length in double stranded deoxyribonucleic acid (DNA), and cleave specific site of the double helix DNA. Each restriction enzyme recognizes a specific sequence of nucleotide bases and cleaves the DNA along the molecule. Bacteria prevent their own DNA from being degraded by methylating their recognition sequences, which sequences are thus modified and protected from the endonucleases. The specific cleavage sites of restriction enzyme are close to the recognition sequences, and therefore restriction enzymes that recognize different nucleotide sequences can be purified from different bacterial species. They function like genetic scissors which allow DNA to be cut at desired sites and therefore become powerful tools in genetic or molecular manipulation.

[0004] The restriction enzyme and its corresponding methylase constitute the restriction-modification system (R-M) of a bacterial species. R-M system in bacteria protects against invasion of foreign DNA. The restriction endonuclease recognizes a specific sequence and the cognate methyltransferase modifies the same sequence to differentiate self-DNA from foreign DNA. Thousands of restriction enzymes have been purified and characterized.

[0005] Because of the abovementioned properties, the use of restriction enzymes are broadly applied in genetic engineering, DNA or gene cloning and gene mapping.

[0006] There are more than 20 putative R-M systems discovered in *H. pylori*

26695 and J99 strains based on sequence homology. Previous studies show that there are 14 Type II R-M systems with biochemical activities in *H. pylori* 26695 strain. The R-M systems of these two strains are very different when the complete sequences of 26695 and J99 strains are compared. The difference of R-M systems results in the barrier of interstrain plasmid DNA transfer and chromosomal DNA transformation. The biological significance of such diverse and complicated R-M systems in *H. pylori* is still unclear.

[0007] Isoschizomer are restriction enzymes that recognize the same sequence. However, the isoschizomers from different sources showed various sensitivities to different modified DNA. Therefore, several restriction enzymes from different sources but cutting the sequence at the same location will be employed together to obtain a better cleavage effect in the target nucleotide sequence.

[0008] The present invention therefore provides a restriction enzyme which can specifically recognize and cut a particular nucleotide sequence in order to provide alternative choices for cleaving DNA in the biotechnological manipulation of genetic engineering and gene cloning, and to improve the cutting efficiency. In addition, it offers a better cleavage effect for target nucleotide sequences which can not be efficiently cut with known restriction endonucleases.

SUMMARY OF THE INVENTION

[0009] The primary object of the present invention is to provide a novel type II restriction endonuclease which recognizes and cuts DNA only at a particular sequence of nucleotides.

[0010] Such novel type II restriction enzyme is obtained from R-M system of a *Helicobacter pylori* strain which is publicly deposited as CCRC17132 in Culture Collection and Research Center of Food Industry Research and Development Institute (Taiwan) on February 12, 2001.

[0011] Using transposon shuttle mutagenesis of *H. pylori* (CCRC17132) allows the identification of different mutant strains. Cell adherence assay is employed thereafter to screen low adherence of these mutant strains. Mutants exhibiting decreased adherences are observed under a microscope. Light microscopic observation reveals a significant elongated morphology, as shown in Fig. 1.

[0012] The DNA sequences for each of the mutants are determined with conventional inverse polymerase chain reaction and sequence analysis. The same locus in these six mutant strains is interrupted by a transposon gene. Nucleotide and amino acid sequences show no homologies with the published sequences of *H. pylori* 26695 and J99 strains.

[0013] This transposon gene comprises a novel open reading frame (ORF) which contains 1617 base pairs (SEQ ID NO : 2) and encodes a peptide of 538 amino acids (SEQ ID NO : 3). The amino acid sequence SEQ ID NO : 3 shares 24% identity with a putative nicking enzyme of *Bacillus halodurans*. In addition, SEQ ID NO : 3 is homologous to the known type II restriction endonucleases PleI and MlyI with 23% and 20% identity respectively.

[0014] The 1617 base pairs of sequence SEQ ID NO : 2 is expressed with

conventional methods of protein expression, and purified through conventional purification methods. The protein obtained is termed HpyC1I, and the number is designated SEQ ID NO:3. The purified protein, HpyC1I, shows endonuclease activity with a non-palindromic recognition sequence of 5'-CCATC-3' (designated SEQ ID NO : 1) and cleaves the fourth base downstream from the recognition sequence of the upper strand and the fifth base from that of the lower strand of SEQ ID NO : 1. The recognition and cleavage site of HpyC1I is identical to those of the known restriction endonuclease BccI respectively after comparison. This result shows that HpyC1I is an isoschizomer of BccI.

[0015] On the other hand, two ORFs are located upstream of the gene encoding HpyC1I after further analysis. HpyC1I and these two putative methyltransferases (M1.HpyC1I and M2.HpyC1I) function together to compose a restriction-modification (R-M) system to protect *H. pylori* CCRC17132 from invasion of foreign DNA.

[0016] The present invention will be further explained in the following embodiment illustration and examples. However, the present invention is not limited to these examples. The present invention may be altered or modified and all such variations are considered within the scope and spirit of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] FIG. 1 shows a comparison of morphology of wild type and mutant strains of *H. pylori* CCRC17132 observed with a light microscope at high magnification (1,000 X) after Gram staining, wherein A is wild type and B is mutant strains.

[0018] FIG. 2 shows adherent ability of wild type and mutant strains of *H. pylori* CCRC17132, wherein the adherent ability is expressed as a percentage in comparison to wild type (taken as 100%). The values are means of triplicate data.

[0019] FIG. 3 shows gene arrangement of R-M system from wild type and mutant strains of *H. pylori* CCRC17132, and the relative areas from *H. pylori* 26695 and J99 strains, wherein arrows represent position and orientation of open reading frame. The figure is illustrated but not according to percentage of a real size. The name and the size of each locus are labeled above the arrows.

[0020] FIG. 4 shows the *lambda* DNA patterns after digestion with HpyC1I and BccI respectively in a 1.5% agarose gel, wherein Lane 1 is uncut *lambda* DNA; Lane 2 is *lambda* DNA digested with HpyC1I; Lane 3 is *lambda* DNA digested with BccI; and Lane M is 1 kb DNA marker.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

Example 1: Screening mutants with adherence assay

[0021] Conventional transposon (mini-TnKm) shuttle mutagenesis is employed in a clinical isolate *H. pylori* CCRC17132 to obtain mutant strains. A total of 1500 *H. pylori* mutant strains are obtained.

[0022] To identify the adherence of mutant strains, 24-well culture plates are used to screen each of the mutant strains in duplicate. SC-M1 used in this study, is a

cell line established from primary human gastric cancer tissue. This cell line is proved to be Le^b negative and sLe^x positive by monoclonal antibodies against Le^b (Seikagaku, Tokyo, Japan) and sLe^x (Chemicon, Temecula, CA) respectively.

[0023] First of all, the SC-M1 cells are grown in RPMI 1640 medium (Gibco BRL, Rockville, MD) supplemented with 10% fetal calf serum (FCS). The cells are plated in 24-well culture plates and grown in a humidified atmosphere at 37°C with 5% CO_2. Infections are performed to the cells by adding *H. pylori* at a multiplicity of infection (MOI) of 100. After 30 minutes of co-cultivation at 37°C, non-adherent bacteria are removed with PBS buffer washing for three times. SC-M1 cells with adherent *H. pylori* are trypsinized, serially diluted in normal saline, and spread on the Columbia blood agar plates. Recovered adherent bacterial colonies are counted. Wild type CCRC17132 strain is served as a positive control, and the adherent ability of each mutant strain is compared to that of the wild type strain.

[0024] Six mutant strains are obtained, which are co-cultivated with SC-M1 cells for 30 min. These six mutant strains reveal a 5 to 10-fold decrease of the recovered adherent bacteria counts compared to those of wild type strain (FIG. 2).

[0025] The morphologies of *H. pylori* wild type and mutant strains are observed with a light microscope after Gram staining and recorded by CoolSnap-pro software (Media Cybernetics, Silver Spring, MD). More than ten fields are examined on each slide and the lengths of bacteria are measured in 30 bacteria of 5-10 different fields by CoolSnap-pro software (Media Cybernetics).

[0026] Light microscopic observation reveals elongation of the mutant strains (FIG. 1). The lengths of wild type strain are 4.3 ± 0.82 μm and those of the mutant

strains are 8.7 ± 1.50 μm in average. Therefore, the decreased adherent ability might due to abnormal appearance or other indirect effects.

Example 2: Identification of inserted gene with inverse PCR and DNA sequencing

[0027] To identify genetic loci interrupted by the transposon, genomic DNA of mutant strains are extracted and subjected to inverse PCR and DNA sequencing analysis. The mini-TnKm insertion site for each of the mutants is determined and compared with the NCBI BLAST databases (http://www.ncbi.nlm.nih.gov/BLAST) as well as the *H. pylori* genome database (http://www.tigr.org).

[0028] Results show that these six mutant strains are interrupted by the mini-TnKm at the same locus. The transposon insertion site of these six mutants is at the 773th nucleotide of this locus. This transposon gene comprises a novel open reading frame (ORF) which contains 1617 base pairs (SEQ ID NO : 2). The nucleotide and amino acid sequences show no homologies with the published sequences of *H. pylori* 26695 and J99 strains. The amino acid sequence SEQ ID NO : 3 encoded by SEQ ID NO : 2 is compared with NCBI BLAST databases. The amino acid sequence shows 24% identity with a putative nicking enzyme in *Bacillus halodurans*, and 23% and 20% identity with two Type II restriction endonucleases PleI and MlyI , respectively.

[0029] There are two ORFs located upstream of SEQ ID NO : 2 after further analysis. Both the upstream ORFs contain a methyltransferase domain, with 780 base pairs and 846 base pairs in length respectively, and are termed *hpyC1IM1* (SEQ ID NO : 4) and *hpyC1IM2* (SEQ ID NO : 5). Based on protein function predictions and gene alignments, the present inventors propose that these 3 ORFs form an operon and

function as an R-M system. This 3.3 kb DNA fragment containing SEQ ID NO : 2, SEQ ID NO : 4 and SEQ ID NO : 5 is absent in both *H. pylori* 26695 and J99 strains (FIG. 3). It has been deposited in DDBJ/EMBL/GenBank with an accession number of AB118944.

Example 3: Expression and purification of restriction endonuclease

[0030] To analyze the activity of this restriction endonuclease, the *hpyC1IR* gene is subcloned into pET28c plasmid and expressed in *E. coli*. Purification of His-tag fusion protein is carried out under Ni-NTA agarose column chromatography.

[0031] In the beginning, the gene encoding HpyC1I SEQ ID NO : 2 is amplified by using PCR and cloned into a pGEM-T easy plasmid (Promega, Madison, WI, USA). The resultant plasmid pGEM-T easy/*hpyC1IR* is then digested with NotI (New England Biolabs, Beverly, MA) and ligated in-frame into pET28c plasmid (Novagen, Darmstadt, Germany). The resulting pET28c/ *hpyC1IR* plasmid is transformed into an *E. coli* strain BL21(DE3). The HpyC1I protein is expressed under 1mM IPTG (isopropyl-β-D-thiogalactoside) induction at room temperature. The His^6 tag protein is purified with a Ni^{2+}-NTA agarose column (Qiagen, Hilden, Germany). The enzyme activity of purified protein is determined on HpyC1I digested *lambda* DNA.

[0032] The endonuclease activity of purified protein HpyC1I (SEQ ID NO : 3) is detected by cleavage of *lambda* DNA. The preferred reaction conditions are under 1× NEB buffer 1 (10 mM Bis Tris Propane-HCl, 10 mM $MgCl_2$, 1 mM DTT pH 7.0) supplemented with 100 μg/ml BSA and incubate at 37℃. About 60 ng purified protein (0.1 μl) can digest 1 μg of *lambda* DNA in one hour at 37℃.

Example 4: Recognition and cleavage site of HpyC1I

[0033] To determine the recognition and cleavage site of HpyC1I, cloning and sequencing of the HpyC1I digestion products from bacteriophage *lambda* DNA (New England Biolabs) are performed. The HpyC1I digested fragments are blunted by T4 DNA polymerase and cloned into the EcoRV (New England Biolabs) site of pBR322 plasmid. Because the EcoRV site of pBR322 is in the tetracycline resistance gene fragment, the AMP^r (ampicillin resistant) and Tc^s(tetracycline sensitive) transformants are selected. Plasmid DNA is isolated from the abovementioned colonies and 10 of the restriction fragment-vector junctions are sequenced.

[0034] Comparisons of the 10 junction sequences indicate that no sequence is the same. Therefore, HpyC1I does not recognize and cut within some sequence fragments. Further analysis identifies a putative non-palindromic recognition sequence in the cloned inserts at a constant distance from the junction. Therefore, HpyC1I belongs to type II restriction endonuclease. The enzyme recognizes a 5 base-pair asymmetric sequence, 5'-CCATC-3' (SEQ ID NO : 1), and cleaves DNA downstream of the recognition site, after nucleotide 4 and 5 in the upper and the lower strand respectively. The double-strand cleavage of HpyC1I produces a one-base 5'-protruding end as shown in Table 1.

[0035] In addition, searches in REBASE database (http://rebase.neb.com) reveal that both the recognition and cleavage sites of HpyC1I are identical to restriction endonuclease BccI. Therefore, HpyC1I is an isoschizomer of BccI. The reaction conditions, R-M genes alignment, and the HpyC1I digestion patterns of *lambda*, pBR322 and phiX174 DNA are all the same with BccI (FIG. 4).

Table 1. Cloning and sequencing of the HpyC1I digestion products from bacteriophage *lambda* DNA are employed to determine the recognition and cleavage site of HpyC1I.

Position in *lambda* DNA	DNA sequence around HpyC1I cleavage site of *lambda* DNA
1325-1364	5'-*CTGGCCAAAGT*CCATC*CGTG*↓**GCTCCA CGCCAAAAGTGAGA**-3'
1596-1635	5'-**GAAAAGACCGGGATCTGGAC**↓*CCGT*GA TGG*CATTCTCTGGT*-3'
4797-4836	5'-*TGCTCGATATGGACACGCCC*↓**GGCGGG ATGGTGGCGGGGGC**-3'
4970-5009	5'-**CGGACAGGCT**CCATC**GGCGT**↓*CATGAT GGCTCACAGTAATT*-3'
9581-9620	5'-*CAGTGGTATGA*CCATC*ACCG*↓**TGAACG GCGTTGCTGCAGGC**-3'
9855-9894	5'-**GTGGAAGACG**CCATC**AGAAC**↓*CGGCG CGTTTCTGGTGGCGA*-3'
11833-11872	5'-*TCCTGCAGGCGGATTACAAC*↓**ACGCTG ATGGCGGCGGCGAA**-3'
12404-12443	5'-**TGAAGACCAGCTTCGCGGGA**↓*ACTG*GA TGG*CAGGCCTGAAG*-3'
39312-39351	5'-*AGACTATCGCA*CCATC*AGCC*↓**AGAAAA CCGAATTTTGCTGG**-3'
39588-39627	5'-**ATCTATGAAAAACATCGCCG**↓*CACA*GA TGG*TTAACTTTGAC*-3''

Bold letters: HpyC1I digested *lambda* DNA after cloning.

Italic letters: Neighboring nucleotides around the HpyC1I digested *lambda* DNA after cloning.

Boxed regions: Recognition site of HpyC1I enzyme.

Vertical arrows: Cleavage site of HpyC1I enzyme.